DOCUMENT RESUME

ED 242 717 SP 024 154

AUTHOR Stiggins, Richard J.; Bridgeford, Nancy J.

TITLE Performance Assessment for Teacher Development.
INSTITUTION Northwest Regional Educational Lab., Portland, OR. Center for Performance Assessment.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Jan 84
CONTRACT 400-83-0005
NOTE 36p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Elementary Secondary Education; Evaluation Methods; *Evaluation Utilization; Faculty Development; *Formative Evaluation; *Job Performance; School Districts; *Summative Evaluation; *Teacher Evaluation; *Teacher Improvement

ABSTRACT

An investigation was made of current uses of formative and summative teacher evaluation. A review of the literature on teacher evaluation systems revealed that most systems attempted to accomplish two potentially conflicting purposes: to encourage teacher development and improve instructional quality (formative), and to judge teacher effectiveness and acquire evidence of incompetence (summative). An analysis of legal constraints and a series of case studies in local districts, illustrated the fact that collective bargaining agreements have been a major force in the school district's definition and limitation of teacher evaluation procedures. This factor has led to emphasis on summative evaluation, with the result that assessment practices contribute relatively little to the identification of teachers' strengths and weaknesses or skill improvement. Information acquired through summative teacher evaluation is seldom used organizationally to plan staff development activities. A review of the results of the case studies produces a list of barriers to formative teacher evaluation. It is concluded that it is both feasible and advisable to emphasize formative evaluation and develop an environment conducive to its success. A set of guidelines is presented for developing an effective formative evaluation system. (Author/JD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED242717

SCOPE OF INTEREST NOTICE
The Eric Facility has assigned this document for processing to: SP EA
In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
M. Margaret Rogers
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

PERFORMANCE ASSESSMENT FOR TEACHER DEVELOPMENT*

Richard J. Stiggins
and
Nancy J. Bridgeford

Center for Performance Assessment
Northwest Regional Educational Laboratory
300 S.W. Sixth Avenue
Portland, Oregon 97204

January 1984

*This research was conducted under contract #400-83-0005 with the National Institute of Education. Opinions expressed in this publication do not necessarily reflect the position of NIE, and no official endorsement should be inferred.

SP024154

Abstract

The purpose of this investigation was to explore current uses of formative and summative teacher evaluation, and to determine the extent that the assessment of performance contributes to improving teachers' skills. The study also examined the characteristics of the assessment methods used to judge teacher performance and effective ways of linking teacher evaluation with teacher development. As a result of a review of teacher evaluation literature, an analysis of legal constraints and a series of case studies in local districts, the authors conclude that assessment practices contribute relatively little to the identification of teachers' strengths and weaknesses or to improving skills. Moreover, information acquired through teacher evaluation is seldom used organizationally to plan staff development activities. Problems with current evaluation measures are specified and potential solutions are proposed.

PERFORMANCE ASSESSMENT FOR TEACHER DEVELOPMENT

Interest in assessing teacher performance has heightened in the last decade. More than ever before, parents want assurance that teachers are competent professionals. Just when schools are plagued by spiraling costs, declining enrollments and dwindling resources, educators are pressed to seek better ways of assuring quality instruction. Effective teacher evaluation is one way to reach that goal.

Educators are concerned about teacher evaluation for many specific reasons. However, as Millman (1981) notes, we can distinguish between two major types of teacher evaluation: formative and summative. The goal of formative evaluation is to identify teachers' own strengths and weaknesses and plan appropriate professional development activities. Summative evaluation results, on the other hand, provide a base for administrative decisions involving hiring and firing, promotion and tenure, assignments and salary.

Clearly, formative and summative evaluations serve different purposes--both important. Summative evaluations are designed to ensure that highly qualified educators enter the profession and continue teaching. Formative evaluations help those already teaching to develop and refine vital skills. Most teacher evaluation conducted today attempts to do both simultaneously. In practice, however, most evaluation practices address summative goals. Formative teacher evaluation--potentially important in instructional improvement and individual development--often assumes a secondary role.

Neither summative nor formative teacher evaluations, as currently conducted, serve their respective purposes as well as they might. Neither the environment in which summative evaluation is conducted nor the assessment procedures used for that evaluation are as effective as they could be. Formative evaluation offers a potential seldom realized simply because it demands more time and effort than many evaluators can afford. In both cases, a new course of action is needed if the potential of teacher evaluation is to be realized. The research project described in this report is an attempt to chart part of that course.

This investigation addresses problems and solutions associated with formative evaluation. We do not wish to imply that formative evaluation is necessarily more important than summative. Both are potentially valuable. But issues of teacher selection, retention and promotion are already receiving widespread consideration; we feel teacher improvement needs more effort and attention. If overall school improvement is our primary goal, then teachers' professional growth and development become paramount.

Further, performance assessment--the observation and rating of behavior--can make a significant contribution to teacher evaluation when used in a formative way. For reasons that are outlined below, the rules governing summative evaluation often preclude the use of much potentially valuable performance data. Formative evaluation, free of such constraints, offers a richer source of performance information on which to base teacher development.

Our goals in this study were first, to understand current teacher evaluation, both its problems and potentials; and second, to identify ways that evaluation can be effectively used to promote teacher

development. To reach these goals, we addressed four questions: How are teachers typically evaluated?, How are evaluations used to foster teacher improvement?, In what ways can evaluation and development be more effectively linked?, and What barriers prevent linkage?

In answering these questions, we (1) reviewed current research on teacher evaluation, (2) reviewed current laws, regulations and contracts pertaining to local evaluations, and (3) conducted case studies of teacher evaluation policies and practices in four school districts. In addition, we conducted a working conference, emphasizing formative teacher evaluation practices and priorities, with task forces of principals, teachers and district administrators from each of the four case study districts.

RESEARCH ON TEACHER EVALUATION

The majority of school districts look to teacher evaluation as a means of improving teaching performance (Educational Research Service, 1978; Bolton, 1983). But given the ineffectiveness of current evaluation approaches, most are not achieving this goal to any impressive degree. To understand why, we explored four dimensions of teacher evaluation: the current context; characteristics of constructive, formative evaluation systems; research on the status of evaluation practices in schools; and the role--actual and potential--of evaluation in fostering teacher growth and development.

The Context of Teacher Evaluation

The current context of teacher evaluation is changing. Donald DuBois, former coordinator of staff development in Salem, Oregon, explains part of the impetus behind this change: "Teacher evaluation, historically, has been a mess. Teachers often feel naked and defenseless by the 'inspection' and 'report card' system. For principals, the teacher evaluation process is a gut wrenching, time consuming duty" (Lewis, 1982).

Educational administrators are aware of problems with current evaluations. Responding to a national survey by the American Association of Secondary Administrators (Lewis, 1982), administrators specified the following needs:

- Better definitions of effective teaching. Although many evaluation programs attempt to define effective teaching, most definitions center on teachers' behaviors--not on appropriately measured outcomes.

- More trust in the process. As one superintendent put it, " We need to know how to evaluate people and get them to feel good about it." In many places, the "spirit" of evaluation has been so structured by teacher contract agreements that it is almost "pro forma."

- Proof of the link between evaluation and instructional improvement. Until there is some specific indication that the process is worth the trouble, some say it will remain "pro forma."

- More specifics on evaluation techniques. Conferences, personal goal-setting, classroom observations--these are common approaches to evaluation, but administrators want to do them better.

- More sensitivity to the needs of the evaluator, primarily the principal. Many participants feel they have neither the skills nor the time for successful evaluations. Evaluators wonder what kind of training they should have and how they should be evaluated to be sure the system works.

As these comments show, administrators are often frustrated by current practices. Evaluation is time consuming, potentially disruptive to staff-administrator relationships, often distrusted and criticized by teachers, and seemingly ineffectual in improving instruction.

Teachers also are critical of evaluation procedures. They often contend that the assessment methods used are inappropriate: the performance criteria by which they are judged are either unspecified or too general; classroom observations are infrequent and superficial; the factors evaluated often have little relationship to instruction; supervisory evaluations are too often subjective, based more on personal characteristics than instructional skill; and results are either not communicated or are not useful in improving performance (Natriello and Dornbush, 1981; Borich and Fenton 1977; Bolton, 1973).

Recent surveys of teachers' attitudes about evaluation bring these problems into sharper focus. For example, they report that only 28 percent of the items in school districts' evaluation checklists examined relate to the instructional role of the teacher. A few additional items deal with such relevant personal characteristics as responsibility and enthusiasm. But the remaining checklist items relate to behavior outside the classroom: memberships in organizations, and participation in professional, social and administrative activities. In other words, the criteria used in these instruments"...seem to be largely unrelated to improving teaching performance and provide little assistance in changing teaching practices."

In studies conducted by Natriello and Dornbusch (1981), teachers noted that they viewed their evaluation systems as generally unsound, overly subjective and unaffected by their efforts. Teachers in these studies indicated that evaluation criteria were seldom shared with teachers, that teachers were uninformed about the information collected to evaluate their performance, and that minimal time was taken to communicate evaluation results to them. The authors note, "The teachers in our interview study reported that on the average they received formal evaluations from their principals once every three years." Moreover, in these teachers' perspectives, evaluations were unrelated to the sanctions or rewards of the system and hence "had little value."

Levin (1979), in a summary of research on teacher evaluation, argues that "research provides little support for current practices in teacher evaluation." He goes on to comment, "One of the few things that can be safely said is that the prevalent system of evaluation...through

observation by supervisors is biased and subjective. The use of techniques that have greater promise for providing objective data, such as observation instruments...is as yet uncommon."

Designing a More Constructive Environment

Dissatisfaction with current evaluation procedures and outcomes has prompted many educators to propose substantive revisions--revisions in planning, in designing performance criteria and acquiring meaningful data, and in communicating results to teachers. The following suggestions represent a concise cross-section of current thought on what strategies might make evaluation practices more constructive and effective.

Manatt (1982), a major proponent of an evaluation model being tried in school districts across the country, advocates an evaluation system manifesting these features: (1) teacher involvement in the evaluation process; (2) centralized and collaborative development of performance criteria based on research and on local priorities; (3) goal setting; (4) multi-dimensional methods for assessing teachers' skills, including objective data gathering, self- and peer evaluation; (5) analysis of results with teachers and development of specific job targets for improvement; and (6) inclusion of a pre-observation conference to acquire background data, and a post-observation conference to mutually analyze classroom data and set goals for improvement.

Manatt's model strongly reflects the positive impact of clinical supervision, a collegial process of professional development designed by Harvard School of Education faculty in the 1960s. Clinical supervision

is a system in which teacher and supervisor work together to set goals and determine progress. This collaborative model includes three major steps: pre-observation conferences to exchange background information and to mutually establish goals and methods for gathering data; multiple classroom observations that include collection of objective data; and post-observation conferences to analyze and verify data and to collaboratively develop a systematic plan for improvement. In addition, teacher and supervisor communicate extensively throughout the evaluation process to ensure that the teacher has a real voice in determining evaluation procedures and setting goals for professional growth. Although clinical supervision methods have been adopted procedurally by districts across the country, the heart of the teacher development process--frequent observation and discussion--has not been successfully incorporated into most local systems.

In response to valid concerns about the perfunctory nature of most evaluations and reliance on subjective observational data, many educators urge the use of (1) assessment methods that give more adequate and objective data about classroom interactions (verbatim records, charts of classroom interactions, records of questioning, or reinforcement strategies); and (2) use of multiple evaluation procedures (student and peer evaluation, assessment of student products) to provide a more comprehensive picture of the teacher's performance (O'Hanlon and Mortensen, 1977). Levin (1979) reinforces the need for "more extensive use of student evaluations and less reliance on ratings by principals and other supervisors." He also concludes that "reliance on a single evaluation technique is unwise" since it reduces the possibility that the teacher will be judged fairly.

Following a comprehensive analysis of current teacher evaluation practices, Darling-Hammond, Wise and Pease (1983) specify four minimal conditions for a successful teacher evaluation system:

- All individuals in the system understand the criteria and processes for teacher evaluation;
- All participants understand how these criteria and processes relate to the basic goals of the organization; i.e., there is a shared sense that the criteria reflect the most important aspects of teaching, that the evaluation system is consonant with their educational goals and conceptions of teaching;
- Teachers perceive that the evaluation procedure enables and motivates them to improve their performance; and principals perceive that it enables them to provide instructional leadership;
- All individuals in the evaluation perceive that the evaluation procedure allows them to strike a balance "between adaptation and adaptability, between stability to handle present demands and flexibility to handle unanticipated demands" (Weick, 1982, p. 674). That is, it allows participants to achieve a balance between control and autonomy.

Achieving a More Constructive Environment

How extensively do current school practices incorporate these commonly advocated criteria? In summarizing a series of surveys of evaluation practices, Knapp (1982) contends that despite strong advocacy of multiple information sources, involvement of students and peers, and more objective means of collecting data, schools have not really changed their approach to teacher evaluation. Principals still do most of the observing; staff are seldom involved in planning; and there is little real effort to use evaluation outcomes in designing constructive staff development.

Similarly, a recent analysis of 32 highly-developed current teacher evaluation systems across the country, completed by the Rand Corporation under sponsorship of the National Institute of Education, provides these insights regarding evaluation practices (McLaughlin, 1982):

Exemplary programs displayed a number of common features. Nearly all of the 32 programs studied required a pre-evaluation conference (88%), classroom observation (100%), post-evaluation conference (100%), a written action plan following evaluation (88%), action plan follow-up (81%), and the participating principal as primary evaluator (78%). Few, however, used self-evaluation (38%), peer evaluation (25%), or students' achievement data (22%) in evaluating teachers.

Despite these similarities, McLaughlin suggests "...there is scant agreement about instrumentation, frequency of evaluation, role of the teacher in the process, or how the information could or should inform other district activities." In other words, little consensus exists about "best practice." Moreover, although 74 percent of the districts named school improvement or staff development as the primary goal of the system, few districts established the necessary links between teacher evaluation and staff development to make that goal achievable. There is, as the study notes, "a general lack of integration between teacher evaluation and staff development or district curriculum guides." Thus, although exemplary programs appear to emphasize staff development and school improvement, teacher evaluation operates, more often than not, as an independent, self-contained system, not an integral component of a broader staff development program.

Effective Formative Evaluation Elusive

As recent studies point out, even highly-developed teacher evaluation programs seem to lack procedures or organizational links essential for systematic individual or staff development through teacher evaluation.

As Knapp (1982) notes, strong formative evaluation programs required both staff involvement and a specified relationship between teacher development and evaluation. But despite the urgings of researchers and educators themselves, not much has happened. Knapp states, although effective evaluation of individual teachers can provide "a more accurate picture of an individual teacher's needs than, for example, the group needs assessments commonly used...systematic evaluation of individual teachers does not as yet appear to be a standard part of staff development planning."

Holley (1982) contends that districts need to make better use of evaluation data. "When evaluator ratings are summarized across competencies or areas, the training needs of both evaluatees and evaluators emerge." District and building summaries can be instrumental in identifying staff development needs. "The data," argues Holley, "should be captured and used for the improvement of both the evaluation process and instruction."

While teacher evaluation practices are becoming more systematic procedurally, most are still inzufficient to support viable teacher improvement programs. Teachers want, at the very least, an evaluation system that provides accurate information on classroom needs, opportunity to acquire and master new learning approaches, and collegial support when instigating needed changes. These activities demand more time,

instructional involvement and more thorough assessment than many principals seem to find manageable. As a consequence. practices become more formalized, remaining basically unchanged. Moreover, as the next section indicates, changes in practice may occur more in response to external pressures than to internal needs.

FACTORS REGULATING TEACHER EVALUATION PRACTICE

State teacher evaluation laws and regulations can influence local evaluation practices, as can collective bargaining agreements. This analysis explores whether these laws and contracts promote or constrain the use of evaluation for teacher development.

We begin the analysis of state laws and regulations with a brief summary of the national picture, then comment on district/teacher association contracts.

Teacher Evaluation Laws

Twenty-six states currently have laws requiring teacher evaluation according to Wuhs and Manatt (1983). Though an equal number currently have no laws, the number of such laws has increased dramatically during the past twelve years, with over 80% of all laws enacted since 1971.

Wuhs and Manatt, for example, reports that in nearly all states improvement is a primary purpose; in almost half, evaluation data are also used for personnel purposes. Beckham (1981), by contrast, reports that less than half of the states list school or teacher improvement as their primary purpose, and that the remainder of laws serve personnel decision making functions. This apparent discrepancy may simply indicate that most evaluation laws serve multiple purposes and often claim to address both formative and summative issues.

Three-quarters of the states leave control of evaluation procedures to local districts. Very few specify criteria to be evaluated. And still fewer provide any guidelines for the development of local systems. So it is apparently local decisions, not state mandates, that determine

most district evaluation procedures. Since local procedures are often negotiated as part of collective bargaining agreements, as a general rule, state laws would appear to have minimal influence on promoting or limiting any local emphasis on formative teacher evaluation. They may, however, indirectly impede formative practices by requiring certain procedures such as use of uniform, mandated evaluation reports, or rigid specification of teaching criteria unrelated to district identified teaching priorities. In such instances, evaluators may first meet the formal, state-mandated requirements and then may decide that implementation of added formative practices is not (a) necessary or (b) possible, given other time demands.

Although some state laws do include some provision for teacher development, most are far less prescriptive than the law passed in Connecticut for the purpose of mandating formative evaluation. Connecticut State Department of Education guidelines specify the following evaluation methods.

- Cooperative planning between professionals and evaluators of the objectives of each individual evaluation, the evaluation procedures, and the process of evaluating the system by staff,
- Clear specification and communication of the evaluation purposes as well as the specific responsibilities and tasks that will serve as the frame of reference for individual evaluations,
- Opportunity for teachers to evaluate themselves in positive and constructive ways, and
- Emphasis on diagnostic rather than evaluative assessment with specific attention given to analyzing difficulties, planning improvements and providing clear, personalized, constructive feedback.

As our discussion will show, these state guidelines attempt to establish a stronger tie between teacher evaluation and teacher development than do most.

The Impact of Collective Bargaining Agreements

A major force in defining and limiting district teacher evaluation procedures has been the collective bargaining agreement. Its chief impact has been to provide due process safeguards for teachers. According to Strike & Bull (1981), who studied numerous such agreements, these contracts affect the evaluation process by: specifying the frequency of evaluations or observations, informing teachers of evaluation criteria, restricting some methods of information gathering (e.g., unannounced visits, secret monitoring, electronic equipment, etc.), and specifying who can and cannot participate in the evaluation process. They also often require formal communication of evaluation results, regulate written reports (e.g., confidentiality, opportunity for a written response by teachers), require remediation for negative evaluations, allow union representation at all conferences and procedures, and necessitate that notice and reasons be filed for disciplinary action, dismissal or demotion.

Overtly, the effect of these contracted agreements is to promote uniformity and specificity in evaluation procedures. But because of the adversarial nature of many contract negotiations, teacher evaluation is becoming more rule governed and "focused more on adherence to clearly specified minimum work standards" (Darling-Hammond, et al., 1983). Collective bargaining has done little to promote links between teacher evaluation and individual development. Although it has often clarified evaluation requirements, it also has had the effect of making evaluation procedures more impersonal and rule governed and unintentionally introduced another dimension of divisiveness into the process.

The legal and contractual factors governing teacher evaluation are typically designed to standardize evaluation requirements and to promote due process in summative evaluation. By promoting adherence to fair practices in personnel decision making, however, they have, in effect, directed attention towards legal requirements and away from methods for promoting teacher growth and development. Although it is encouraging to find some state laws attempting to strike a balance between formative and summative teacher evaluation, the laws only set an overall context for practice in districts and schools. Their effectiveness still depends on interpretation and implementation. Regulations may indeed enhance the potential for improvement, or they may be viewed as just one more requirement to be met, thus having little substantive impact on instructional changes.

DISTRICT TEACHER EVALUATION PRACTICES

What impact do existing laws and contracts actually have on local evaluation practices? Do the same teacher evaluation issues identified in national studies also concern local districts? What practices do local programs use to promote teacher development? To answer these questions, we asked four Pacific Northwest school districts to participate in case studies of their teacher evaluation systems. The results added much detail to our understanding of the evaluation environment.

Case Study Description

To learn about teacher evaluation practices in these districts, we conducted 17 interviews with district administrators in charge of teacher evaluation and with building principals (elementary, junior high, and high school). In addition, each principal was asked to name four teachers from her or his school who would respond to a questionnaire on teacher evaluation practices. Forty-eight teachers received questionnaires, and thirty-six responded.

Participants were asked to describe teacher evaluation practices from their viewpoint. Further, they were asked if and how they used results to plan teacher development. Interviews and questionnaires touched on state and district policies, development of evaluation procedures, methods for gathering information on teacher performance, methods for communicating evaluation results to teachers, and relative satisfaction with their evaluation system. In outlining ways evaluation data are used, administrators and principals described the specific

decisions--summative and formative--influenced by the results. They were also asked what role teacher evaluation plays in promoting teacher growth and improvement. Results of the interviews and questionnaires are summarized below.

Case Study Results

In summarizing case study results, we found evaluation procedures to be strikingly similar across districts. For example, all districts used a three-stage evaluation process including a pre- and post-observation conference, and classroom observation. The observations, conducted either by principals or vice-principals, were the central feature of all evaluations. They occurred formally either once or twice a year. Peers and students were seldom involved in the evaluation; self-evaluations were cursory if done at all, and student achievement scores were not used. Results of the evaluator's observation were communicated both in person and in writing to the teacher. The written reports called for supervisors to describe teachers' strengths and weaknesses on either state or locally specified criteria. None of the districts used rating scales or indications of performance levels to identify teachers' skills. Finally, training provided evaluators ranged, in one district, from a frequent, integrated program that involved all staff to intermittent or sporadic training in other districts. In addition, both teachers and administrators saw room for improvement in the evaluation process and made specific recommendations about needed changes.

Teachers' Perspectives on Needed Changes. The primary goal of our case studies was to identify barriers precluding use of teacher

evaluation results for teacher growth and development. In the questionnaire, teachers were asked for their perspectives on

- Needed changes in the teacher's role in evaluation,
- Needed changes in district procedures, and
- Ideas for improving the quality of teaching in the district.

With respect to the teacher's role, over half (53%) spontaneously urged more opportunity for collegial observation, and for self-evaluation through goal setting and videotaping. Others suggested giving teachers more knowledge about what constitutes effective teaching, and more proficiency in evaluating lessons.

Recommendations for improvements in the overall evaluation system were far more extensive. Repeatedly, teachers suggested more frequent, formal and informal observations, greater use of peer observation and self-evaluation, and more effective preparation and training for evaluators. In addition, they called for better observational strategies, more effective communication of results, with emphasis on specific suggestions for improvement, increased skill among evaluators, and better general management of evaluation. Teachers also noted that they needed quality inservice training to improve their skills.

Teachers seemed to agree that to be effective, evaluation must (1) be a valued schoolwide priority, not just a requirement; (2) occur frequently enough so that outcomes reflect actual classroom activities; (3) incorporate methods that provide relevant, specific, and complete information; and (4) involve evaluators trained to provide specific, useful suggestions for improvement. Repeatedly teachers called for more

frequent feedback, and for constructive criticism, not "vague generalities that hide mediocre teaching."

Administrators' Perspectives on Needed Changes. How satisfied were principals and district staff with their evaluation systems? Responses differed considerably across districts.

In two districts, administrators were generally satisfied with the evaluation process, but concerned with the amount of time available to conduct observations. In the two other case study districts, administrators were less satisfied. Reasons for dissatisfaction included teachers' lack of trust in the evaluation process, lack of clarity in criteria, and the fact that evaluation seemed more oriented to meeting state standards than promoting improvements.

Similarly, not all agreed that teacher evaluation led to teacher improvement. Some felt that the goal setting process was a major step toward improvement, and that increasingly, evaluation was focusing more on improvement than on judgment. Others felt the link was weak at best. Principals also generally acknowledged that evaluation results were not directly used to plan school or districtwide staff development, and that local supervisors did not include evaluation results in setting instructional priorities. Evaluation was, however, used by some evaluators to help teachers identify individual goals and to specify a plan of action for the year. The completion of these plans and their effect on instruction was seldom monitored.

These administrators cited four major barriers limiting the development of a more formative evaluation system: (1) teachers' lack of trust in the process; (2) insufficient time for evaluation; (3) the

adversarial context of evaluation; and (4) principals' skills as evaluators. The trust and time issues were most frequently mentioned. Principals also noted these problems: inconsistent evaluation procedures, imprecise requirements, lack of training for evaluators, imprecise dismissal procedures, overly general evaluation outcomes, teachers' resistance to change and inadequate staff development. Administrators noted too that many principals did not know how to establish evaluation programs or set realistic priorities.

When asked how evaluation could be more directly related to the improvement of teaching, administrators recommended changes in system management, including increased staff involvement in goal setting and emphasis on improvement as a district priority, improved methods of conducting observations, more time allowed for evaluation and observations, development of evaluators' skills, a stronger link between evaluation and staff development, and accountability for all principals conducting evaluations.

Yes, say supervisors, evaluation could be more effective in diagnosing teachers' needs and improving their skills--but only if the process changes in many ways. And the major changes called for closely parallel those concerns identified in national studies of teacher evaluation.

A Conference on Teacher Evaluation

As a conclusion to the case studies, we invited teams of educators--each team including a district administrator, principal and teacher--from each of the four case study districts to attend a working conference and consider, in greater depth, methods for more effectively

linking evaluation and staff development within the districts. Conference teams discussed two major issues: (1) the barriers to more formative evaluation, and (2) potential ways of linking evaluation more closely to teacher improvement.

After reviewing the results of the case studies, conference participants worked together to produce the following common list of barriers to formative teacher evaluation (listed in order of importance) in their districts:

1. Evaluators often lack important skills needed to evaluate, and the training needed to solve this problem is frequently not available, not used, or ineffective. At least two sets of skills are lacking:

   a. skills in evaluating teacher performance, and

   b. skills in communicating with teachers about the evaluation process and results.

2. There is often insufficient time for both evaluation and followup. A continuous cycle of feedback and growth is needed to promote teacher development. The competing demands of education frequently push evaluation to a low priority status.

3. The process(es) for linking staff development and teacher evaluation are not clear. We lack a clear goal for formative teacher evaluation (i.e., an image of the desired system) and a plan for achieving that goal. State laws and district policies and procedures do not reflect that goal; and individuals (teachers and administrators alike) within the system have yet to provide the support needed to make evaluation results truly productive. Despite an important emphasis on protecting the due process rights of teachers, evaluation systems lack a similar commitment to promoting professional development.

4. Trust in the evaluation system is often lacking among educators functioning within that system. Unclear or unacceptable performance criteria combined with lack of teacher involvement in developing performance criteria and infrequent and superficial observations, tend to breed skepticism among teachers about the value of results. The adversarial relationship between districts and collective bargaining units also breeds distrust.

Having identified a list of significant barriers, conference participants then turned to the task of finding solutions. Simple solutions were not expected nor sought. Nevertheless, participants generated a list of starting points from which to begin various assaults on key barriers.

Individuals noted that commitment to effective teacher evaluation as an important means to school improvement must be given priority status within the district. In addition, district staff need to determine the foremost purpose of their evaluation--either staff improvement or personnel decision making--and develop procedures appropriate to accomplishing that purpose. Added to these initial steps, conference participants called for:

- Involvement of all staff in the planning process,
- Identification of meaningful and relevant performance criteria as the basis of the evaluation,
- Evaluators trained to pinpoint teachers' skills in need of development,
- Inclusion of other sources of information about teachers' proficiency including data from peers, self and students,
- Development of a comprehensive staff development program for evaluators as well as teachers, and
- Adequate resources--time and money--to develop a thorough program of feedback and development for staff.

IN SEARCH OF SOLUTIONS

Teacher evaluation, as the case studies and summaries of national teacher evaluation practices indicate, is viewed as an important school-based method for improving teachers' skills. In practice, however, evaluation has substantially less impact than is desired. Despite increasing emphasis on improving the quality of teacher evaluation programs, most "improvements" seem directed at systematizing procedures. As a result, regulations abound. Most teacher evaluation systems now require regular annual or biannual evaluations, specify the general performance criteria to be used, define the procedures in the evaluation cycle, call for written documentaion of results, and require that those results be formally reviewed with teachers. Some teacher evaluation programs have increased teachers' participation in evaluation by including personal goal setting and carefully outlining plans for accomplishing goals. But educators generally concur that even highly developed evaluation systems are not helping teachers either individually or collectively to improve their skills.

Some of the changes needed to make evaluation more effective--from teachers' and supervisors' perspectives--as mentioned are these: teachers and administrators need to collaborate on the goals, criteria and procedures for evaluation; evaluation procedures need to be geared to diagnosing teachers' skills and aiding them in making instructional changes; supervisors need more time, support, and training to diagnose instructional problems; and evaluation methods need to be more soundly based and linked to rewards and sanctions within the organization.

Persistent problems suggest that a number of substantial changes are needed in the organization of evaluation systems, the emphasis given instructional improvment, the kind and quality of information provided teachers, and in the ways evaluation is conducted.

Making the Purpose Clear

Most teacher evaluation systems have attempted to simultaneously accomplish two potentially conflicting purposes: to encourage teacher development and improve instructional quality (formative), and to judge teacher effectiveness and acquire evidence of incompetence (summative). As a result, most have succeeded only marginally, doing whatever necessary to meet minimal legal requirements. In practice, summative evaluation has usually taken precedence; formative activities, by nature more time consuming and demanding, has been dealt with superficially if at all. As a consequence, evaluation systems have not provided adequate diagnosis and assistance to support teacher improvement. Further, the dual purpose of most evaluation programs has increased tension, in many instances undermining the trust, honesty, openness and motivation needed to promote experimentation with new teaching approaches. For formative evaluation to work most effectively, it needs first to be specified as a primary purpose of evaluation, and second to include appropriate procedures for acquiring information.

In the authors' opinion, it is both feasible and advisable to emphasize formative evaluation and to develop an environment conducive to its success. Although this may be accomplished in many ways, success of the activity appears to hinge on a number of important steps. First, it

demands an openness to change and commitment to improvement. Teachers and supervisors must agree on priorities, sharing decisions on what needs to be accomplished and when. Second, it involves cooperatively selecting training opportunities, determining the degree to which objectives are accomplished and new skills learned (selecting criteria for performance), incorporating multiple sources of information to determine the effect of new teaching approaches on students, and sharing resources to support changing behaviors. In effect, success will require that teachers and administrators work together cooperatively, as mutually supportive allies.

Emphasizing teacher development as a major purpose of evaluation requires strategies different from those commonly used in summative evaluation--as the following paragraphs show.

Improving the Quality and Availability of Information

Responsibility for summative evaluation falls most frequently to the school's principal or vice-principal. More often than not, once-a-year observation is the sole basis for determining teachers' performance and identifying needed skills. In formative evaluation, numerous information sources may be tapped. Peers, students, and teachers themselves offer a broad spectrum of perspectives, thus increasing the odds that strengths and needs will be accurately identified. Further, responsibility for formative evaluation can be placed first and foremost in the hands of each individual teacher and employ relevant, useful data from sources--like students and fellow teachers--who are thoroughly familiar with the classroom environment.

Many other information resources, usually disregarded in summative evaluation, are appropriate in formative evaluation. For example, student achievement data can be useful. Although standardized achievement test data are considered an inappropriate basis for summative evaluation, teachers' own day-to-day classroom measures provide diagnostic information that may be a vital component of a teacher's self-evaluation of instructional strengths and weaknesses.

Ensuring Adequate Performance Criteria

The significance of relevant performance criteria can scarcely be overstated. Criteria present stumbling blocks to sound assessment if and when they:

- Focus on personal characteristics rather than instructional skills.
- Call for inferences about teaching behavior that compromise reliability.
- Are too general to provide diagnostic information.
- Are unclear or unrelated to professional practices of teachers.

Formative teacher evaluation can be more effective if these factors are considered in establishing the performance criteria that key the process.

First, criteria should relate to student outcomes defined as important by current research and be identified collaboratively by teachers and principals. The emphasis should be on behaviors that seem to make a difference, such as "the clarity of a teacher's presentations." Although researchers acknowledge that not all behavior works in all settings, there is growing evidence now that certain instructional methods, such as those associated with direct instruction, have impact on student achievement in many contexts.

Second, each performance criterion should describe some teacher behavior or characteristic of the classroom environment that can be consistently evaluated--regardless of when the evaluation occurs or who observes the behavior. As Millman (1982) notes, "How a teacher is judged often depends on who does the observation and on what day it is performed."

Third, each criterion should be clear and specific enough to yield diagnostic information. Ratings that do not suggest how performance might be improved are of little value to a teacher.

Fourth, each performance factor must be endorsed by the teacher as appropriate for his or her classroom. To merit such endorsement, criteria must be (1) valid within the unique learning environment established by the teacher, (2) appropriate for the content taught and the instructional method(s) used by that teacher, and (3) flexible enough to allow the teacher a choice of instructional strategies.

Fifth, performance criteria need to be practical as well as relevant. Though criteria need to be diagnostic, a long list of minutely specific behaviors that cannot be rated, communicated or addressed in a reasonable time is likely to generate confusing feedback and planning problems. Some balance is needed between diagnostic precision and cumbersome detail. If the list of important performance criteria grows excessive, evaluators should set priorities and address only part of the list in a given term, semester, or year.

Sixth, in reviewing criteria, the teacher and evaluator should identify relatively weak areas of performance and mutually design steps to improve those areas. Moreover, in all formative evaluations, it is

inappropriate to compare one teacher's performance with that of another for the purpose of ranking teachers by proficiency. The uniqueness of learning environments, student groups, instructional styles and teacher groups make such comparisons meaningless. In addition, using teacher norms or ranks invariably promotes a defensiveness that is counterproductive. After all, professional development, not criticism for its own sake, is the whole point of the system.

Finally, all desirable change depends on establishment of effective channels of communication between teacher and supervisor. For many teachers, evaluation results have not been communicated either constructively or diagnostically. As one educator commented, "fault finding without suggestions for remedy, categorizations (e.g., good, average) that provide little diagnostic assistance, generalities that appear to have little factual basis and reports that make no clear contribution to organizational goals" are not effective forms of communication (Bolton, 1973). This is true, of course, for either formative or summative evaluation.

Useful evaluation results promote instructional awareness and prompt change. To do so, presentation of feedback should occur in an atmosphere of mutual problem solving and trust; teachers need evidence that their efforts toward professional growth will be rewarded with personal recognition and support. As Darling-Hammond, et al. (1983) state, "Effectively changing the behavior of another person requires enlisting the cooperation and motivation of that person, in addition to providing guidance on the steps needed for improvement to occur."

Effective Formative Evaluation

Assessing teachers' performance is an important task. It is hoped that the foregoing discussion will promote better understanding of effective evaluation practice, while at the same time encouraging educators to attempt alternative, dynamic approaches to the formative evaluation process. Those alternatives can work to improve instruction if the following guidelines are observed:

- Select Methods to Match Evaluation Purpose. The purpose of an evaluation system must be clearly identified and understood in order to select appropriate methods. The same procedures cannot constructively and simultaneously serve the needs of those interested in promoting teacher development and those responsible for personnel decisions. Both sets of needs are important. But different methods are needed to address different purposes.

- Involve Teachers in Evaluation. Teachers should be involved in all phases of developing and operating formative systems. Any evaluation program that does not reflect the interests, concerns, aspirations and needs of teachers is doomed to failure. By the same token, teachers must have constructive attitudes to make the system work. Teaching must be regarded as a skill to be learned and participants must be willing managers of their own development, ready to consider, explore, and practice new teaching skills.

- Provide Relevant Training. All evaluators and staff must be thoroughly trained. Everyone involved in the evaluation should know how to use evaluation instruments to acquire useful, objective data, interpret results, and use those results to advantage. Similarly, evaluators should be trained to provide feedback to teachers which is clear, precise and sufficiently diagnostic to promote realistic plans for improvement.

- Increase Sources of Evaluation Data. Thorough formative evaluation should include the perspectives of students, peers, teachers themselves, and supervisors and incorporate several kinds of observation not just once-a-year classroom spot checks.

- Use Meaningful Criteria. Performance criteria must be relevant to desired student outcomes, specific enough to be useful in planning professional development, and accepted as important by each teacher to whom they will apply.

- Relate Results to Organizational Goals. Evaluation results should be used by both teachers and staff development planners to set training priorities and to evaluate success in achieving organizational and personal goals. Successful evaluation is clearly tied to organizational planning. Moreover, the system itself should be evaluated regularly before any procedures become so firmly entrenched that they are unresponsive to change.

The entire framework of this proposal for formative evaluation rests on one overriding assumption: That is, school managers and teachers alike function best within an environment characterized by mutual support, by respect and concern for personal growth, and for the well-being of staff and students. Where such an environment exists, formative teacher evaluation offers great potential for helping teachers learn to teach better.

Given current economic conditions and declining enrollments, fewer new teachers are entering the profession. Therefore, improving the quality of instruction demands developing the skills of teachers already in the classroom. Formative evaluation--a system inherently sensitive to teachers' own needs and goals--can be a vital step in strengthening instructional effectiveness nationwide.

REFERENCES

Beckham, J.C. Legal Aspects of Teacher Evaluation. Topeka, KS: National Organization on Legal Problems in Education, 1981.

Bellon, Jerry J. Teacher evaluation: From the teacher's perspective. CEDR, Phi Delta Kappa, Winter, 1982, Vol. 15, No. 4.

Bolton, Dale L. Realistic methods of evaluating competent teachers. Olympia, WA: School Information and Research Service, May-June, 1983.

Bolton, Dale L. Selection and Evaluation of Teachers. Berkeley, CA: McCutchan, 1973.

Borich, G.D. and K.S. Fenton. The Appraisal of Teaching: Concepts and Processes. Reading, MA: Addison-Wesley, 1977.

Darling-Hammond, Linda, Wise, Arthur E. and Pease, Sara R. Teacher Evaluation in the Organizational Context: A Review of the Literature. Washington, DC: National Institute of Education, 1983.

Educational Research Service, Evaluating Teacher Performance, Arlington, Virginia, 1978.

Holley, Freda, Personnel education: Essentials for success. CEDR, Phi Delta Kappa, Winter, 1982, Vol. 15, No. 4.

Knapp, Michael S. Toward the Study of Teacher Evaluation as an Organizational Process: A Review of Current Research and Practice. Menlo Park, CA: SRI International, March, 1982.

Levin, Benjamin. Teacher evaluation--a review of research. Educational Leadership, 1979, 37(3), 240-245.

Lewis, A.C. Evaluating Educational Personnel. Arlington, VA: American Association of School Administrators, 1982.

Manatt, Richard. Teacher Performance Evaluation - Practical application of Research. Ames, IA: Iowa State University, Occasional Paper 82-1, January, 1982.

McLaughlin, M.W. A Preliminary Investigation of Teacher Evaluation Practices. Santa Monica, CA: The Rand Corporation, 1982.

Millman, Jason. (Ed.). Handbook of Teacher Evaluation. Beverly Hills, CA: Sage Publications, 1981.

National School Public Relations Association. Evaluating Teachers for Professional Growth: Current Trends in School Policy and Programs. Arlington, VA: Autumn, 1974.

0859e

Natriello, Gary and Sanford M. Dornbusch. Pitfalls in the Evaluation of Teachers by Principals. Chicago, IL: University of Chicago, 29(6), 1980-81.

O'Hanlon, J. and L. Mortensen. Approaches to teacher evaluation. CEDR, Phi Delta Kappa, Winter 1977, Vol. 10, No. 4.

Peterson, K. and D. Kauchank. Teacher Evaluation: Perspectives, Practices and Promises. Salt Lake City, UT: Center for Educational Practice, Graduate School of Education, University of Utah, 1981.

Ryan, Doris W. and Edward S. Hickcox. Redefining Teacher Evaluation. Toronto, Ontario: The Ontario Institute for Studies in Education, 1980.

Strike, K., and Bull, B. Fairness and the legal context of teacher evaluation. In J. Millman (Ed.), Handbook of Teacher Evaluation. Beverly Hills, CA: Sage Publications, 1981.

Weick, K.E. Administering education in loosely coupled schools. Phi Delta Kappan, 1982, 673-676.

Wood, Carolyn J. and Paul A. Pohland. Teacher Evaluation: Organizational Maintenance Versus Stimulation of Improved Teaching Performance. Department of Educational Administration, College of Education, University of New Mexico, 1979.

Wuhs, Susan K. and Richard P. Manatt. The pace of mandated teacher evaluation picks up. The American School Board Journal, 1983, 170(4), 28.